УДК 519.68

Є.Н.Манн, Б.М.Голуб, Г.Г.Цегелик, Ю.М.Щербина

МОДЕЛІ ОПТИМІЗАЦІЇ ПАРАМЕТРІВ ОБЧИСЛЮВАЛЬНОЇ СІТКИ

Розглянемо моделі визначення оптимальної довжини інформаційного пакета при пересилці даних між вузлами обчислювальної сітки і оптимального розподілу файлів за вузлами сітки з кільцевою топологією.

Нехай дані зі зовнішнього запам"ятовуючого пристрою /ЗЗП/ ЕОМ$_I$ потрібно передати на ЗЗП ЕОМ$_2$. Повідомлення /записи/ можуть мати велику довжину, тому розбиваються ЕОМ$_I$ на пакети фіксованої довжини в l байт, які на приймальній стороні /ЕОМ$_2$/ знову збираються у ціле повідомлення. Відзначимо, що при розбиранні повідомлення останній пакет може виявитись коротшим.

Найпоширенішою реалізацією протокола інформаційного процесу у випадку, який розглядається, є протокол KERMIT . Формат пакета протокола KERMIT містить службову інформацію /заголовок, довжина пакета, номер, тип, контрольне поле/ та інформаційне поле даних. Позначимо довжину службової інформації через l_0 /байт/, а поле даних - через l .

При виникненні у системі передачі даних спотворень останні повинні бути одразу виявлені. З цією метою у контрольне поле записують контрольну суму, використовуючи який-небудь код з виявленням похибок. Нехай β - середнє число логічних операцій на I байт при формуванні контрольної суми. У випадку сумування і перевірки контрольної суми спеціальним пристроєм, який характеризується пропускною здатністю /ПЗ/, параметр β буде дорівнювати нулю.

Сформований пакет передається на канал зв"язку /КЗ/. Під КЗ будемо розуміти функціонально зв"язані між собою технічні засоби /кодери, модеми, підсилювачі тощо/, які мають характеристикою ПЗ, і середовище передачі. Під пропускною здатністю КЗ розуміємо ПЗ "вузького місця" КЗ, тобто найменш продуктивного пристрою або середовища передачі. Позначимо ПЗ канала через С /байт/сек/.

У процесі перетворень і передачі сигналів в КЗ вони можуть спотворюватись. Відісланий пакет зберігається у буферному запам"ятовуючому пристрої у вигляді копії доти, доки від приймального вузла не надійде підтвердження про прийом і відсутність помилок передачі.

У випадку від"ємної відповіді здійснюється повторне відсилання пакета. Повторне відсилання здійснюється також, якщо протягом деякого часу t_T інформація про прийом відсутня. Час t_T називається часом тайм-аута.

Передача інформації може здійснюватись у різних режимах залежно від будови лінії зв"язку: напівдуплексної або дуплексної. У випадку використання напівдуплексної лінії витрачається час для перемикання з одного напрямку на інший. Позначимо цей час через θ. У схемі передачі даних, яка розглядається, дуплексну лінію можна ототожнити з напівдуплексною при $\theta = 0$.

Введемо наступні позначення:

ℓ_1 – довжина пакета-підтвердження /байт/;

p – ймовірність спотворення байта /кодової посилки/ в КЗ;

τ – мінімально можливий час з моменту передачі пакета довжиною ℓ байт до прийому пакета-підтвердження $\Delta = t_T - \tau$;

t_n – очікуваний час достовірної передачі пакета довжиною $\ell + \ell_0$ байт.

У співвідношенні

$$t_n = \frac{\tau + \Delta(1 - q^{\ell_1})}{q^{\ell + \ell_0 + \ell_1}} \qquad /I/$$

$q = (1-p)^K$, причому $K = 8$ при синхронному способі передачі і $K = 10$ або $K = 11$ при асинхронному способі /залежно від кількості стопових бітів/.

Нехай пересилання інформації здійснюється відповідно із описаною схемою. Припустимо, що зв"язок між ЕОМ встановлено та створення сигналів можливі лише в КЗ.

Позначимо через t_0', t_0'' – середній час пошуку файла на ЗЗП відповідно ЕОМ$_I$ та ЕОМ$_2$; S_0', S_0'' – швидкість обміну інформації між ЗЗП і процесором ЕОМ$_I$ та ЕОМ$_2$ відповідно /байт/сек/; v_1, v_2 – швидкодія процесорів ЕОМ$_I$ та ЕОМ$_2$ /логічних операцій/сек/; α – середня кількість операцій на I байт поля даних при формуванні пакета; γ – середня кількість операцій на I байт службової інформації при формуванні пакета; L – середня/очікувана/ довжина повідомлень /запису/ у байтах.

Будемо вважати, що всі введені параметри фіксовані, окрім довжини поля даних ℓ.

Час $\tau(\ell)$ можемо подати у вигляді

$$\tau(\ell) = K\cdot(\ell + \ell_0 + \ell_1)/c + [\beta\cdot(\ell + \ell_0) + \gamma\,\ell_1]/v_2 + 2\theta,$$

де K = I для синхронного способу передачі і K = I,25 або K = I,375 для асинхронного способу.

Підставляючи $\tau(l)$ в /I/, отримаємо

$$t_n = \frac{R_1 l + R_2}{q^l},$$

де

$$R_1 = (K/c + \beta/v_2)/q^{(l_c + l_1)},$$

$$R_2 = [K(l_c + l_1)/c + \beta l_c + \gamma l_1)/v_2 + 2\theta + \Delta(1 - q^{l_1})]/q^{l_c + l_1}.$$

Нехай $x > 0$. Позначимо через $I(x)$ цілу частину числа x. Враховуючи час формування пакета та обміну інформацією з ЗЗП, отримаємо очікуваний час пересилання запису /повідомлення/:

$$T(l) = (t_c' + L/S_o') + I\left(\frac{L}{l}\right)[t_n(l) + t_1(l)] + $$
$$+ m[t_n(\bar{l}) + t_1(\bar{l})] + (t_o'' + L/S_o''),$$

де $t_1(l) = (\alpha l + \gamma l_c)/v_1$; $\bar{l} = L - l \cdot I(L/l)$,

$$m \begin{cases} 0, \text{ якщо } \bar{l} = 0, \\ I, \text{ якщо } \bar{l} \neq 0. \end{cases}$$

Досліджуючи функцію $T(l)$, доходимо висновку, що величина l може істотно впливати на час пересилання запису. Тому доцільно відзначити довжину l поля даних пакета за критерієм мінімуму очікуваного часу пересилання запису:

$$\frac{\min}{l} \{T(l) : I \leq l \leq L, \quad l - \text{ціле}\}. \quad /2/$$

Задачу /2/ можна досить ефективно розв"язувати шляхом неповного перебору допустимих значень l, який базується на твердженні: якщо при переході від l до $l+1$ кількість пакетів не змінюється, то $T(l) < T(l+1)$ і здійснювати обчислення при $l+1$ немає сенсу.

Розглянемо тепер побудову моделі оптимального розміщення файлів за вузлами обчислювальної сітки з кільцевою топологією. У ролі критерію оптимальності приймемо загальну суму добутків об"ємів даних, що становлять запити та відповіді, породжені функціонуванням системи протягом одиниці часу на віддалі, на які ці дані пересилаються. При цьому, якщо вузли сітки пронумерувати у порядку їх наступності по кільцю в сітці, то віддаль між s -м та j -м вузлами визначається як $d_{sj} = \min(|s-j|, n-|s-j|)$, де n - кількість вузлів сітки.

Нехай n /як це вже було відзначено/ - кількість вузлів сітки; m - кількість незалежних файлів розподіленої бази даних;

K_j – j-й вузол сітки; F_i – i-й файл; L_i – об"єм i-го файла; b_j – об"єм пам"яті вузла K_j, яка призначена для розміщення файлів; d_{sj} – віддаль між вузлами K_s та K_j / $d_{ss} = 0$, $s = 1,2,\ldots, n$ /; λ_{ij} – інтенсивність запитів до файла F_i, які ініційовані у вузлі K_j; α_{ij} – об"єм запиту до файла F_i, який ініційований на терміналі вузла K_j; β_{ij} – об"єм запитів даних при виконанні запиту до файла F_i, який надійшов на термінал вузла K_j;
x_{ij} / $i = 1,2,\ldots, m$; $j = 1,2,\ldots, n$ / – величини, що визначаються за формулою

$$x_{ij} = \begin{cases} 1, & \text{якщо файл } F_i \text{ знаходиться у вузлі } K_j, \\ 0, & \text{у протилежному випадку.} \end{cases}$$

Тоді математична модель задачі буде наступною: потрібно знайти мінімум функції

$$L = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{s=1}^{n} \lambda_{ij} (\alpha_{ij} + \beta_{ij}) d_{sj} (1 - x_{ij}) x_{is} \qquad /3/$$

при обмеженнях

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad / i = 1,2,\ldots, m /; \qquad /4/$$

$$\sum_{i=1}^{m} L_i x_{ij} \le b_j \qquad / j = 1,2,\ldots, n /; \qquad /5/$$

$$x_{ij} \in \{0,1\} \qquad / i = 1,2,\ldots, m;\ j = 1,2,\ldots, n /. \qquad /6/$$

Для розв"язування задачі /3/–/6/ розроблено алгоритм, який складається із визначення початкового розподілу файлів і перерозподілу файлів у випадку, коли початковий не є оптимальним. При цьому початковий розподіл файлів завжди оптимальний, якщо на об"єм пам"яті кожного вузла, яка призначена для збереження файлів, не накладається умова /5/.

Для визначення початкового розподілу будується матриця $\Theta = (\Theta_{is})_{m,n}$, де

$$Q_{is} = \sum_{j=1}^{n} \lambda_{ij} (\alpha_{ij} + \beta_{ij}) d_{js} \text{ —}$$

сума добутків об"ємів даних, які пересилаються із вузла K_s при функціонуванні системи протягом одиниці часу на віддалі, що на них ці дані пересилаються, у випадку зберігання файла F_i у вузлі K_s. Після цього для всіх $i = 1,2,\ldots, m$ знаходиться $\min\limits_{1 \le s \le n} Q_{is}$. Нехай

$$\min_{1 \le s \le n} Q_{is} = Q_{is_i} \qquad / i = 1,2,\ldots, m /,$$

тоді початковий розподіл $X = \{x_{ij}\}_{m,n}$ визначається наступним чином $x_{is_i} = 1$ / $i = 1,2,\ldots,m$ /,

$x_{ij} = 0$ / $i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$; $(i,j) \neq (i,s_i)$ /.

У тому випадку, коли початковий розподіл не є оптимальним, здійснюється перерозподіл файлів, суть якого полягає у наступному.

Для кожного вузла K_j перевіряється умова /5/. Якщо для деякого індекса j ця умова не виконується, то із переповненого вузла K_j здійснюється перерозподіл файлів. При цьому на кожному кроці із вузла K_j завжди перерозподіляється той файл, якому відповідає мінімальне збільшення значення цільової функції. Якщо розподіл файлів із вузла K_j завершений, то цей вузол вважається закритим для перерозподілу файлів, тобто до нього неможна більше поміщати інших файлів, і здійснюється перехід до перерозподілу файлів із іншого переповненого вузла /якщо такий існує/.

Відзначимо, що при побудові математичної моделі схема обробки запитів не використовувалась. Отже, вона може бути обрана довільно.

Стаття надійшла до редколегії 30.09.90

УДК 518:517.948

М.В.Жук

ЗАСТОСУВАННЯ МЕТОДУ КАНТОРОВИЧА ДЛЯ НЕЛІНІЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ

Розглянемо нелінійне диференціальне рівняння

$$Au \equiv -\frac{\partial p(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y})}{\partial x} - \frac{\partial q(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y})}{\partial y} + r(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}) = f(x,u) \qquad /1/$$

при нелінійній однорідній крайовій умові

$$R[u]\big|_\Gamma \equiv \Big[p(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y})\cos(\nu,x) + q(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y})\cos(\nu,y) + \sigma(s)u\Big]_\Gamma = 0 \qquad /2/$$

де Γ - межа області D, обмеженої по x прямими $x = a$ і $x = b$, а по y достатньо гладкими кривими $y = g(x)$ і $y = h(x)$, причому $g(x) < h(x)$, ν - зовнішня нормаль до Γ.

Відносно заданих функцій припускаємо, що $f(x,y) \in H = L_2(D)$ з нормою $\|f\|^2 = \iint\limits_D f^2(x,y)\,dx\,dy$; $p(x,y,s,t,z)$, $q(x,y,s,t,z)$, $r(x,y,s,t,z)$